

**Stutzman, P.E.**

## **Reference Materials and Proficiency Testing for Validating X-Ray Powder Diffraction Measurements of Cement Composition**

### **1. Introduction**

Estimates on the phase content of cement for process control and specification compliance using the Bogue calculations have been used for the past 75 years. The Bogue method is an indirect approach based upon a bulk chemical analysis and assumptions of the chemistry of the four principal phases. Bias inherent in the calculations resulting from assumptions of composition and from analytical uncertainties have limited their use in accurately characterizing mineralogical compositions [1,2]. X-ray powder diffraction analysis has become more popular within the industry, providing a direct analytical method for a more complete and potentially more accurate estimate of phase compositions of clinker and cement.

Qualitative and quantitative analysis by X-ray powder diffraction is a complicated analytical procedure, particularly for labs not well acquainted with the analysis procedures, cements, or both. The need for a standardized analytical procedure and some means to validate individual laboratory performance was recognized by interlaboratory studies [3,4,5,6] and became a goal of the ASTM C01 Committee on Cement.

In 1978, ASTM C1.23 Subcommittee on compositional analysis initiated a task group (TG) on X-ray powder diffraction (XRD) to develop a test method to improve the precision and bias [3,4,5,6]. Initial work used sets of calibration mixtures, calibration curves, and compounded test mixtures using an internal standard-based analysis [7,8,9]. Peak intensity measurements remained a source of inconsistency in test results between laboratories [10] leading some to conclude that XRD results were not much better than the Bogue estimates [3,6].

The application of the Rietveld method to XRD analysis of cements [11,12] has dispensed the need to develop complicated calibration curves for each phase, using crystal structure models and the entire diffraction pattern to estimate individual phase intensities. While this has made the measurements more consistent within and between laboratories [13], the inherent data normalization in the Rietveld method can pose a hidden problem. This constraint will obscure bias, if present, and apportion bias across all phases relative to their concentration. So, while this approach has successfully addressed the challenges of reference standards and phase intensity measurement, it remains possible to get a suitable-appearing refinement with a highly biased result. Standard test methods and test validation remain necessary to provide assurance of a suitable analysis.

This paper will present a discussion on the development of a standard test procedure using X-ray powder diffraction for cements, the development and role of certified reference clinkers for phase analysis and, proficiency testing for validating laboratory performance that may encompass a wider range of materials.

## 2. Standard Reference Material (SRM) Clinkers

The initial interlaboratory studies by the ASTM XRD TG involved sets of compounded mixtures of lab-synthesized phases to serve as calibration mixtures and a second set to serve as validation mixtures. A longer-term goal of the ASTM TG was to develop a well-characterized set of industrial clinkers to facilitate the test development. These reference materials (RM) were first made available in the mid-1980's [14] with reference values based on light microscopy point count analysis. Subsequently, an additional data set by X-ray powder diffraction analysis provided the second independent analytical method to establish certified reference values and uncertainties [15].

Three SRM clinkers, SRM 2686, 2687, and 2688 were developed as certified reference materials. They exhibit a range of textures and phase abundance and are intended to be used in developing and testing quantitative methods for cement analysis, primarily microscopy and X-ray powder diffraction. SRM2686a (the successor to SRM 2686) has an intermediate grain size, heterogeneous phase distribution and periclase disseminated throughout. SRM2687a has a fine-grained, moderately heterogeneous texture. SRM2688 is coarse-grained with uniformly-distributed silicates and matrix (Figure 1).

## 3. Test Method Development

The initial set of interlaboratory trials for quantitative analysis of the interstitial phases (aluminates, ferrite, and periclase) [7,8] used compounded test mixtures and later the Reference Clinkers to assess precision and bias. Using selective extractions to concentrate the interstitial phases, precisions of 1.1 % within-lab  $1\sigma$  (repeatability) and 1.7 % between-lab  $1\sigma$  (reproducibility) were calculated on a whole clinker basis. The mean values from the participants closely mirrored the known values, leading to a conclusion of no apparent bias.

The efforts of the task group culminated in the first standard test method for X-ray powder diffraction analysis of cements, ASTM C 1365, Standard Test Method for Determination of the Proportion of Phases in Portland Cement and Portland-Cement Clinker Using X-ray Powder Diffraction Analysis [16,17]. ASTM C1365 is a performance-based standard. While most standard test methods use an explicit protocol, this test method allows flexibility in lab procedure but requires that a set of performance criteria must be demonstrated via analysis of certified reference materials.

The test method was initially based upon the internal standard method using calibration curves but the performance-based approach easily accommodated growing popularity of the Rietveld method. The adoption of the Rietveld method however has presented one new shortcoming. Inherent in the calculations is a data normalization that may obscure bias. In the traditional calibration curve approach that assesses each phase individually, having a data set sum to less than or greater than 100 % indicates some problem with the data or the calibrations that needs additional investigation. While the Rietveld approach has improved consistency, differences between laboratory procedures may result in significant disagreement, so the need to demonstrate appropriate test performance remains with the SRM clinkers and proficiency testing using cements providing a means to qualify laboratory performance.

#### 4. Measurement Uncertainty in X-ray Powder Diffraction Measurements of Cements

XRD measurements of hydraulic cements, like any measurement process, are subject to a random and a lab-specific systematic bias (error) [18]. Four factors contribute to testing variability: 1) the operator, 2) the equipment, 3) instrument calibration, and 4) the testing environment [19]. For example, a new operator may not have as much experience in phase identification and quantitative analysis, one instrument may be lower-powered than others, counting statistics (count time, detector design), specimen grinding and mounting procedures, structure models, and analysis codes. While not an exhaustive list, these variables are relatively constant within a lab, imparting some measurement variability between specimen replicates. However, they will generally vary between labs, resulting in larger differences in between lab test results.

Precision is defined within ASTM as “the closeness of agreement among test results obtained under prescribed conditions” [19]. Precision is expressed as repeatability (within-laboratory, designated  $s_r$ ) which excludes the four factors and reproducibility (between-laboratory, designated  $s_R$ ), which includes the four factors. Both measures are expressed as standard deviations of replicate analyses. From the  $s_r$  and  $s_R$ , the maximum difference ( $d2s$ ) between two test results may be calculated by multiplying the respective standard deviation by 2.77 to obtain the 95 % repeatability ( $r$ ) and reproducibility ( $R$ ) limits [19].

These measures form the basis for qualification in ASTM C1365 using a certified reference material portland cement clinker. The qualification provides both the lab and the user of the test results with a degree of confidence that the test can be successfully accomplished. This is particularly important now, where as mentioned earlier, the normalization in the Rietveld calculation may obscure bias.

An alternate validation approach, proficiency testing, provides a means to test method and lab performance using a wider variety of materials. For cements analyses, laboratories provide a single analysis on each of two cements using a consensus for phase abundance and comparison to the ASTM C1365 precision qualification criteria. Both of these approaches will be illustrated.

In the first example, NIST SRM clinker 2688 has been analyzed by 11 individual laboratories using X-ray powder diffraction. Plots of their results are shown in Figure 2 with the three replicate determinations for each phase. The blue box here encompasses the  $\pm 1\sigma$  reproducibility limits for information while the green box represents the qualification criteria,  $\pm d2s$ , the reproducibility limits. The displacement between individual replicate determinations for each lab represents the repeatability and the displacement of the triplicate clusters between labs represents reproducibility. While not illustrated in the plots to maintain simplicity, the  $d2s$  limits for repeatability are coincidentally approximately the width of the blue box. Most labs exhibit a similar repeatability, which generally falls within method limits.

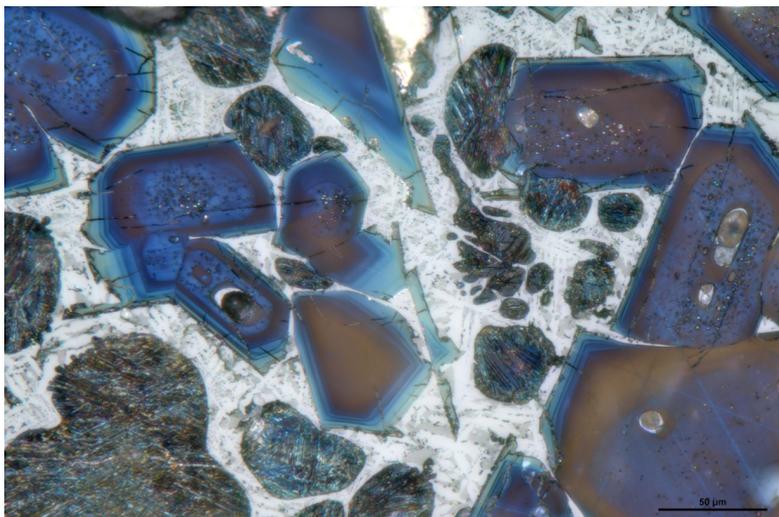
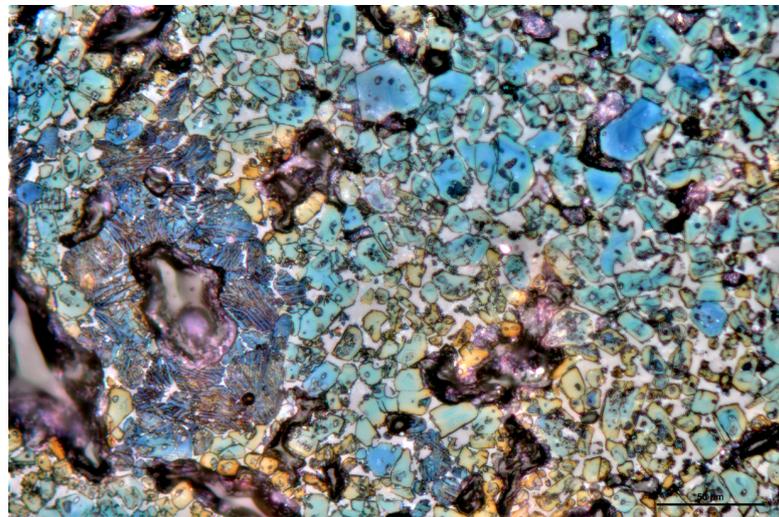
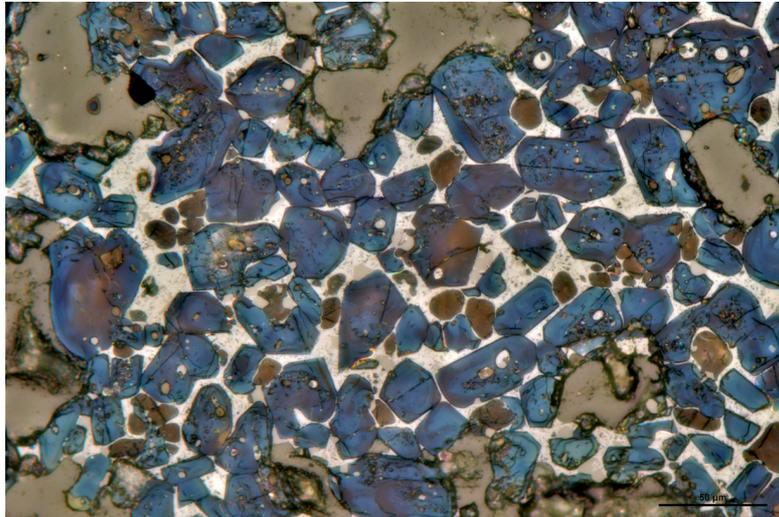


Figure 1. Standard Reference Material (SRM) Clinkers 2686a (top), 2687a (middle) and 2688 exhibit a range of textures and compositions and are used for developing and testing methods for quantitative phase analysis.

Greater differences are seen between lab determinations and, in some cases, one may speculate as to the sources of lab bias. In the cases of labs 6 and 11, for example, there appears to be a correlation between the alite and belite test results. These results may reflect preferred orientation of the silicates and the difficulty in estimating silicate scale factors after orientation corrections have been applied.

So, for a lab to qualify using this SRM, their triplicate results (repeatability) should not be displaced by more than the width of the blue box and their test result must remain within the green box. While both labs 6 and 11 fall within the  $d_2s$  bounds for alite, they fail due to the bias in belite determination. In some cases where lab repeatability is relatively large, bias may be reduced by establishing a test result as a mean of replicate measurements. This can be imagined in the case where lab 1 has its first aluminate determination on the limit and the third exceeding the limit. By re-defining a test result as the mean of  $n=3$  replicates, they would qualify. By doing so, labs 1, 3, and 5 would be able to qualify their ferrite determinations. Alternatively, they could assess their test procedures to determine and resolve the source of the bias.

While the SRM clinkers have been very useful both in test method development and for validating lab performance, there is a need for reference materials that cover the more complicated phase assemblage of cements. While certified reference cements for phase abundance could be developed, proficiency testing may be a more practical approach.

#### **4. Application of Proficiency Testing to Assess Laboratory Bias**

Proficiency testing is a development based upon interlaboratory programs to evaluate test methods and laboratories using paired samples for chemical analyses of rocks and cements by W.J. Youden [20,21,22]. This approach has been used to assess test precision and to evaluate laboratory performance.

Samples of two different, but similar, cements are distributed to a number of laboratories. Each laboratory reports a single test result for each cement phase identified for each of the two cements. A consensus value for each analyte of each cement is taken as the median value of the data. The median is used because of its insensitivity to outliers. Consensus values often mirror the reference values when a large number of labs provide estimates. In the case of alite for SRM 2688 discussed earlier, the certified value of 66.1 % is closely approached at 65.1 % using the median of the eleven labs.

In the following examples, instead of estimating test precision from the data, we apply existing ASTM performance criteria with the Youden graphical display. This allows a visual assessment of the individual and collective test results to assess compliance of a lab, and to provide some insight on systematic and random error in individual test results.

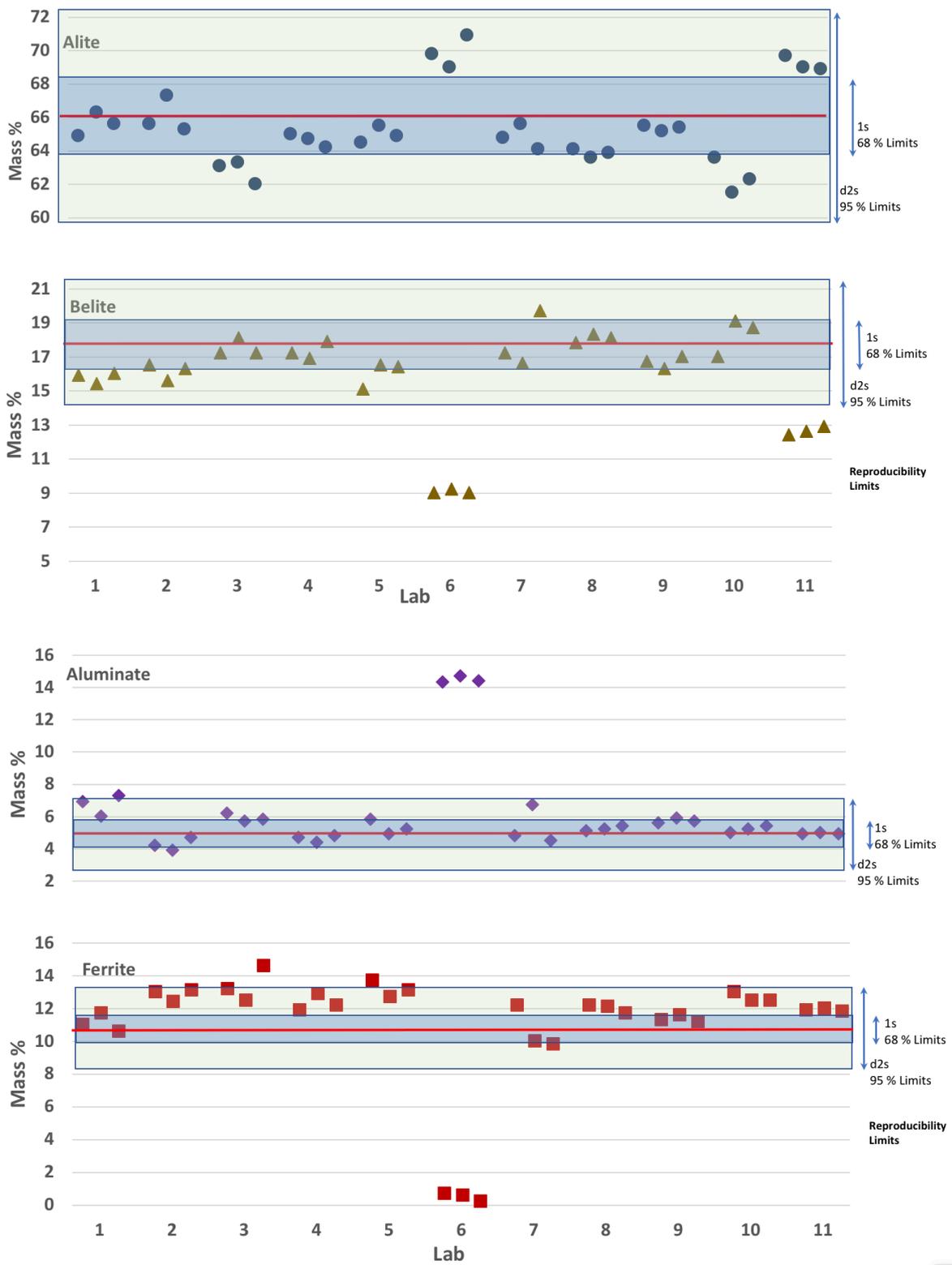


Figure 2. SRM 2688 measurement results from 11 laboratories for alite, belite, aluminate, and ferrite with SRM certificate value indicated by a red line and reproducibility qualification limit boxes for the 95 % limits in green and 1σ in blue.

In the following cement example, a test result is defined as the mean of triplicate determinations. Key assumptions are that repeatability precision from lab to lab is similar and each lab's systematic bias (error) affects both of their test results similarly. A Youden scatter plot is prepared where each point represents results of an analyte from the pair of cements. A consensus value is established by taking the median of the data, as the median is less sensitive to outlying results, and plotted as intersecting lines at right angles to each other.

If the test method is not biased, the points will form a roughly circular pattern centered on the intersection of the median lines. Systematic bias is seen as points falling along the 45 degree diagonal away from the median intersection. Random bias is seen in the degree to which points fall off the 45 degree diagonal line, as points falling in the upper-left and lower-right quadrants have opposite signs to their individual estimates. ASTM E2489 provides a means to calculate test precision using paired sample or single sample data [22].

Using the ASTM C1365 qualification criteria and the Youden plot format, we get Figures 3 through 5. These plots have been augmented with graphical features to denote test method qualification criteria. First, the median lines and 45 degree line segment lengths represent the  $\pm d2s$ , or 95 % limits for reproducibility. The circles represent  $\pm 1\sigma$  and  $\pm 2\sigma$  from the median consensus values for user information since to qualify, lab test results must be within the  $\pm d2s$  limits. An alternate approach to qualify labs was proposed by Youden [21], with lab performance based upon their distance from the consensus as  $\pm 1\sigma$ ,  $\pm 1.5\sigma$ ,  $\pm 2\sigma$ ,  $\pm 2.5\sigma$ ,  $>2.5\sigma$  being assigned values of 4 through 0, respectively, along with a sign of + or - if their bias was high or low. For multi-analyte materials like cements, the results are averaged to establish their overall performance score.

## 5. Observations on the Youden Plots

As an anonymous test program, each lab is aware of their identifying code but not that of any other lab. Plotting of the collective results allows one to quickly compare their reported results against the others and the consensus. The location of their x-y point for each phase reveals both systematic and random error of their testing procedures. Recall that the current reproducibility  $d2s$  limits are indicated by the lengths of the median and 45 degree segments. The inner circle ( $\pm 1\sigma$ ) encloses the best-performing labs and the outer circle the well-performing labs ( $\pm 2\sigma$ ). Labs 3 and 9 exhibited the greatest difficulties in their analyses whereas labs 2 and 22 generally performed within the  $\pm 2\sigma$  level.

For the alite plot in Figure 3, lab 9 (dark triangle) falls outside the  $d2s$  limits, but roughly along the 45 degree line indicating a strong positive bias. Labs 7 (tan triangle) and 10 (blue diamond) exhibit a moderate negative and moderate positive bias, respectively however, they remain within the  $d2s$  limits. In contrast, lab 6 (brown +) is located in the lower left quadrant, reflecting a random error. Lab 6 appears to have had difficulty with Cement A, with a low estimate of alite for Cement B, close to the consensus. The apparent correlation between alite and belite estimates are reflected in the Belite Youden plot where lab 7 is positively biased and lab 10 negatively biased. In addition, labs 4 and 12 show a large systematic bias while lab 20 shows a large random bias. Given alite has a propensity to preferentially orient and the diffraction patterns of alite and belite are quite similar, some labs may be having difficulties with

these effects on the determination of appropriate scale factors that are part of the quantitative estimates.

Laboratory 9 also had significant negative systematic bias for aluminate and ferrite while labs 3 and 11 show significant positive aluminate bias and significant negative ferrite bias for lab 3 and a larger random bias with a low ferrite estimate for cement A. Laboratory 3 also exhibits a large positive bias for periclase as seen in Figure 4.

The calcite plot in Figure 5 shows the elongation of data points reflecting a greater propensity for systematic bias in its phase estimates. This may stem from the propensity for preferred orientation for calcite or from its strong overlap with the alite diffraction pattern. Labs 9 and 3 have systematically high estimates while labs 10 and 12 are systematically low. Lab 5 estimated the calcite content of Cement A close to the consensus but either forgot to include calcite for cement B or had a large random bias component in its test procedure.

In some cases a cement phase is present in only one of the two cements. In these cases, a jittered dot diagram may be used to depict the results along with the consensus value (Figure 6). Labs 9 and 3 performed quite well in their estimation of gypsum, falling well within the  $d2s$  values. Labs 12, 10, and 17 significantly underestimate gypsum relative to the consensus while lab 5 significantly overestimates gypsum concentration. Gypsum is another phase that is very prone to preferred orientation, which is likely to affect its phase estimates. These outlying results suggest that these lab's specimen preparation methods and any use of orientation corrections merit investigation.

The ASTM C1365 qualification performance criteria are based on a more exhaustive inter-laboratory study involving eleven competent labs and four cements compounded from the SRM clinkers. Performing within the  $d2s$  limits or better ensures that a lab is performing at the level comparable to those in the original study. These data should be revised over time as labs become better acquainted with quantitative X-ray diffraction analysis of cements and as instrumentation and processing software improve. Given estimates of repeatability and reproducibility standard deviations can be made from the paired sample and the single sample study, these performance criteria can evolve [22]. The current qualification criteria are on an individual phase basis. The scoring system proposed in [21] would be useful in routine proficiency testing where a running average of all test results would allow labs the ability to examine their test procedures without being excluded for one bad test result.

Zero-value data create some problems for proficiency testing if repeatability and reproducibility values are being estimated. In this example labs did not designate if a zero value was the result of an actual zero determination or that they did not include that phase in their analysis. The test procedure needs evaluation criteria to establish the presence or absence of a phase so it is clear that the value is a zero value rather than a phase that was not identified and therefore left out of the analysis. Selective extraction for the non-silicate phases, an initial qualitative analysis, or requiring a visual check of the refined phase patterns against the cement pattern are all useful methods to clearly establish the presence or absence of a phase. Finally, the effects of preferred orientation on phase estimates appears to be necessary, and recommendations on specimen preparation and the application or orientation corrections should be established.

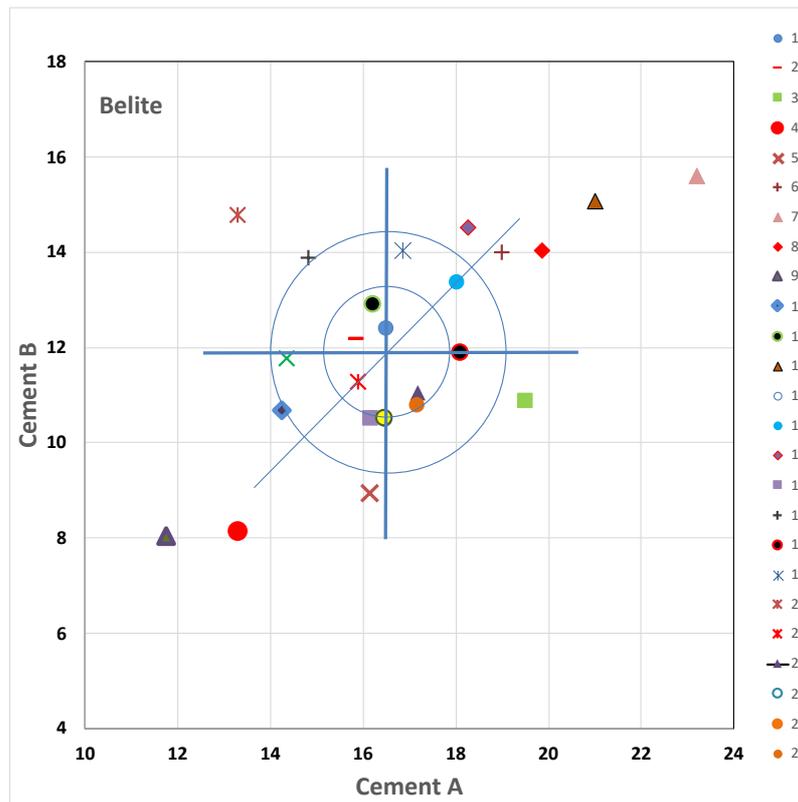
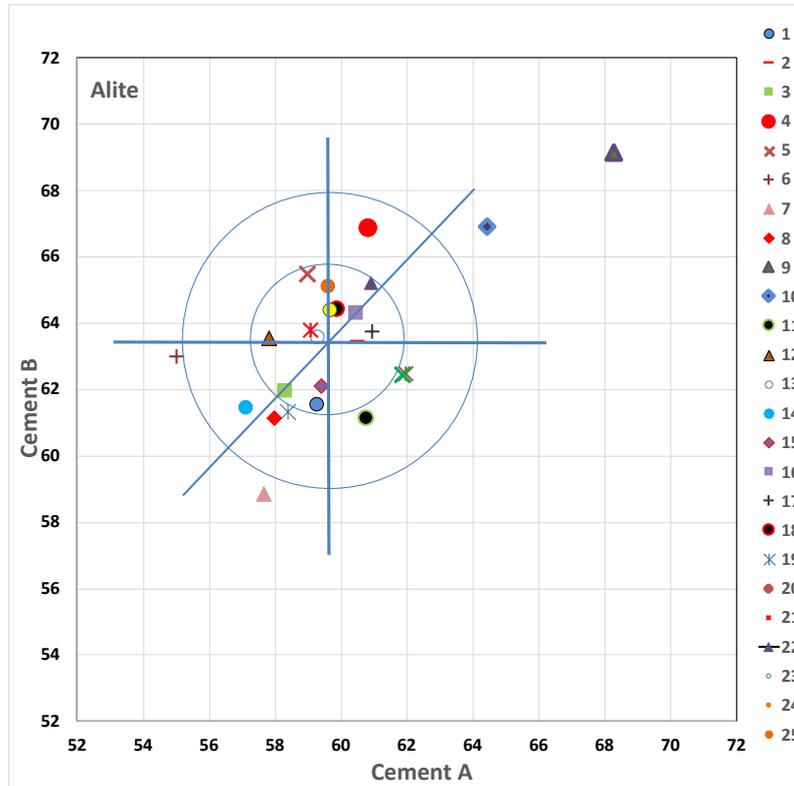


Figure 3. Youden plots for alite and belite provide graphical feedback on lab performance for both systematic and random bias.

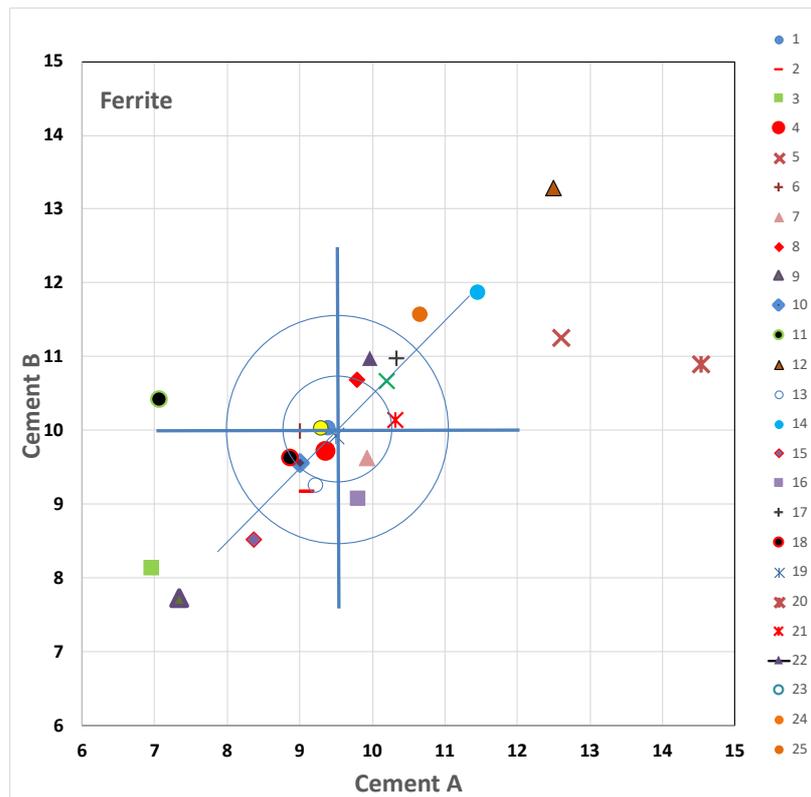
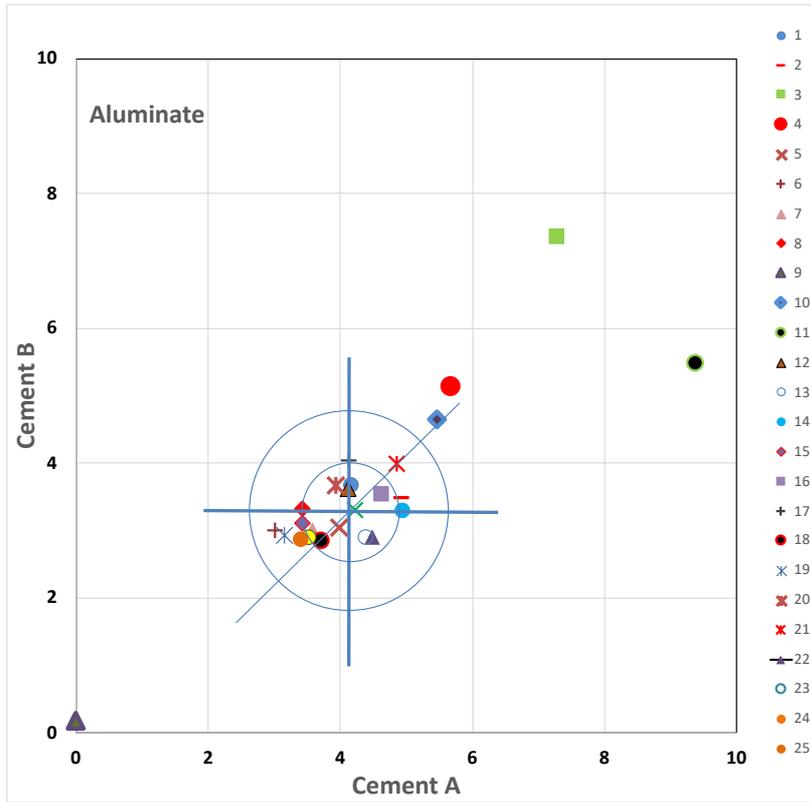


Figure 4. Youden plots for aluminate and ferrite provide graphical feedback on lab performance for both systematic and random bias.

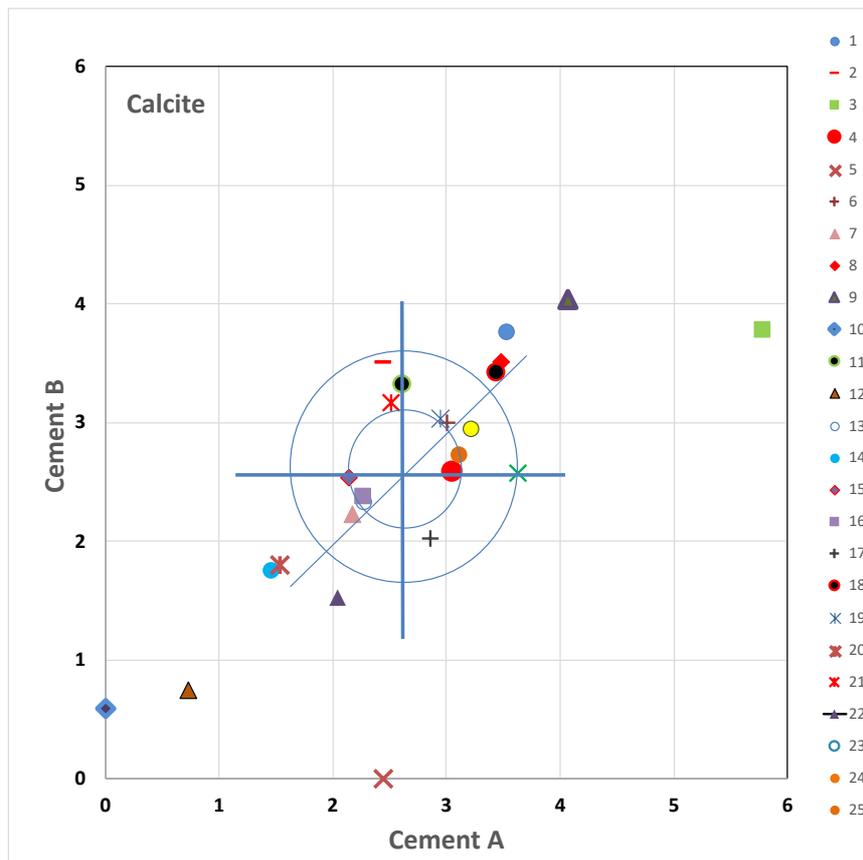
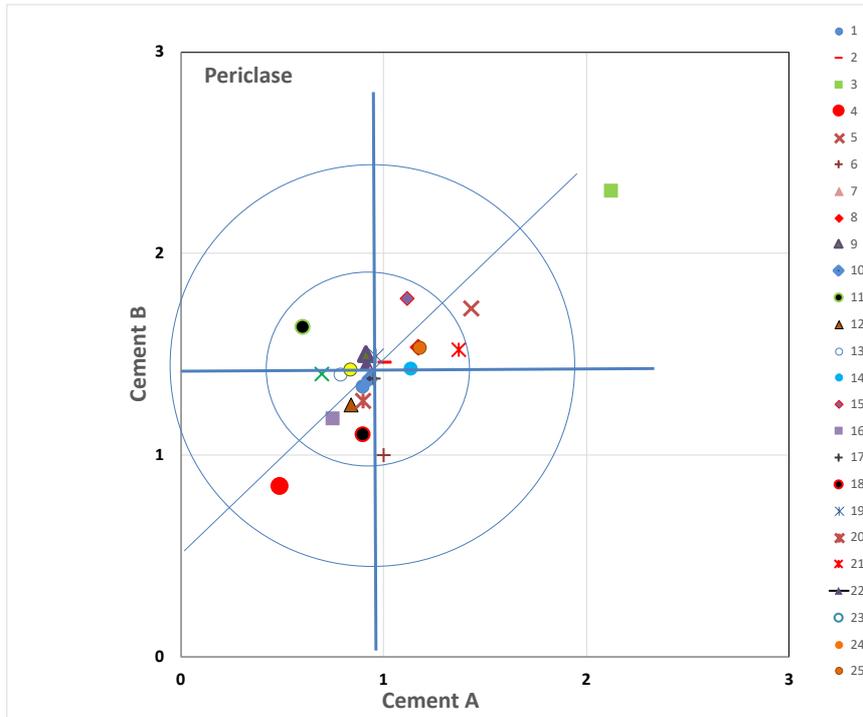


Figure 5. Youden plots for alite and belite provide graphical feedback on lab performance for both systematic and random bias.

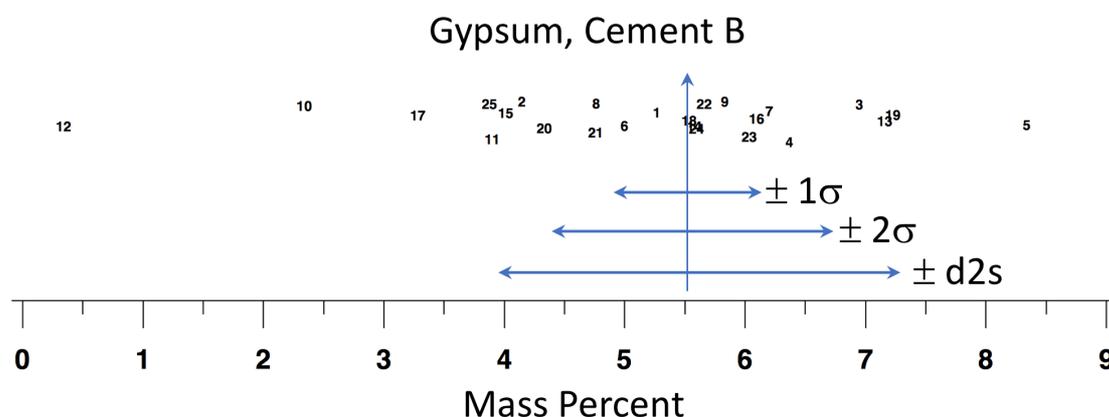


Figure 6. A jittered dot diagram with lab numbers used in place of dots provides a rapid means to illustrate individual lab performance against the consensus for a phase that occurs in only one of the two cements.

## 6. Conclusions

Validation of quantitative X-ray powder diffraction measurements traditionally involved compounding test mixtures of known phase compositions. Validation is particularly important with the Rietveld method as the normalization will obscure bias. ASTM C1365, Standard Test Method for Determination of the Proportion of Phases in Portland Cement and Portland-Cement Clinker Using X-Ray Powder Diffraction, employs a laboratory qualification using replicate determinations of certified reference materials, such as the NIST SRM clinkers. The performance metrics for repeatability and reproducibility precision and for bias in ASTM C1365 are based on multiple inter-laboratory studies and standardized procedures for estimating precision. An alternate validation approach can be found in a modified routine proficiency testing, where a group of laboratories provide a single analysis on each of two cements using the ASTM C1365 qualification criteria. Quantitative feedback provided in graphical format allow participants to assess results for random and systematic error. Proficiency testing of this sort provides participating laboratories with a quantitative assessment of their performance relative to peers using a wider range of materials encompassing the broad spectrum of modern hydraulic cement production.

## References

- <sup>1</sup> H.F.W. Taylor, *Cement Chemistry*, 2<sup>nd</sup> ed., Thomas Telford, London, 1997.
- <sup>2</sup> P. Stutzman, A. Heckert, A. Tebbe and S. Leigh, *Uncertainty in Bogue-calculated phase composition of hydraulic cements*, *Cem. Conc. Res.*, 61 (2014) 40-48.
- <sup>3</sup> A.E. Moore, *Comparison of the Results Obtained for the Compound Composition of Portland Cements by X-Ray Diffraction, Microscopy, and Wet Chemical Methods*, *Analysis of Calcareous Materials*, S.C.I. Monograph No. 18, pp. 372-390, 1964.
- <sup>4</sup> A.E. Moore, *Mineralogical Analysis of Portland Cement by X-Ray Diffraction*, *Proc. Soc. Chem. Ind.*, 18-20, 1970.
- <sup>5</sup> L.P. Aldridge, *International Cement Analysis Study. Part 3. Quantitative X-ray Diffraction Analysis of Cements*, Report No. C.D. 2201, Department of Scientific and Industrial Research, New Zealand, 1975, 20 pp.

- 
- <sup>6</sup> L.P. Aldridge, *International Cement Analysis Study. Part 4: Comparison of Results*, Report No. C.D. 2267, Department of Scientific and Industrial Research, New Zealand, 1978, 18 pp.
- <sup>7</sup> Struble and H. Kanare, *Cooperative Calibration and Analysis of Cement Clinker Phases*, NIST IR 89-4164, 34 pp., 1989
- <sup>8</sup> L.J. Struble, *Cooperative Calibration and Analysis of Cement Clinker Phases, Report 2*, Civil Engineering Studies, Structural Research Series No. 556, University of Illinois, 21 pp., 1990
- <sup>9</sup> L.J. Struble, *Quantitative Phase Analysis of Clinker Using X-Ray Diffraction*, Cement, Concrete, and Aggregates, Winter, 1991, pp. 97-102.
- <sup>10</sup> G. Yamaguchi and S. Takagi, *The Analysis of Portland Cement Clinker*, Principal paper I-3, Fifth International Symposium on the Chemistry of Cement, Tokyo, pp. 181-216.
- <sup>11</sup> H. Möller, *Standardless quantitative phase analysis of Portland cement clinkers*, World Cement, September, 1995, pp. 75-84.
- <sup>12</sup> J.C. Taylor and L.P. Aldridge, *Full-profile Rietveld quantitative XRD analysis of Portland cement: Standard XRD profiles for the major phase tricalcium silicate (C3S: 3CaOSiO<sub>2</sub>)*, Powder Diffraction 8 (3), September, 1993. pp. 138-144.
- <sup>13</sup> P. Stutzman and S. Leigh, *Phase Analysis of Hydraulic Cements by X-Ray Powder Diffraction: Precision, Bias, and Qualification*, Journal of ASTM International, Vol. 4, No. 5, 11 pp., 2007
- <sup>14</sup> H.M. Kanare, *Production of Portland Cement Clinker Phase Abundance Standard Reference Materials*, Final Report, Construction Technology Laboratories, December, 1987, 32 p.
- <sup>15</sup> P. Stutzman and S. Leigh, *Phase Composition of the NIST Reference Clinkers by Optical Microscopy and X-ray Powder Diffraction*, NIST Technical Note 1441, 50 pp., 2002
- <sup>16</sup> ASTM C 1365-06, *Standard Test Method for Determination of the Proportion of Phases in Portland Cement and Portland-Cement Clinker Using X-Ray Powder Diffraction Analysis*, ASTM Book of Standards, Vol. 4.01, ASTM International, West Conshohocken, PA.
- <sup>17</sup> P. Stutzman and S. Leigh, *Phase Analysis of Hydraulic Cements by X-Ray Powder Diffraction: Precision, Bias, and Qualification*, Journal of ASTM International, Vol. 4, No. 5, paper ID JA1101085, 11 pp., 2007
- <sup>18</sup> W.J. Youden, *Statistical Techniques for Collaborative Tests*, Association of Official Analytical Chemists, Inc., 1973 64 pp.
- <sup>19</sup> ASTM E 691, *Standard Practice for Conducting an Interlaboratory Study to Determine the Precision of a Test Method*, in Annual Book of ASTM Standards, Vol. 4.01, 2001. ASTM International, West Conshohocken, PA
- <sup>20</sup> Youden, *Evaluation of Chemical Analyses on Two Rocks*, Technometrics, Vol. 1, No. 4., November, 1959, pp. 409-417.
- <sup>21</sup> W.J. Youden, *Statistical Aspects of the Cement Testing Program*, Proc. ASTM, Vol. 59, 1959, pp. 1120-1128.
- <sup>22</sup> ASTM E2489, *Standard Practice for Statistical Analysis of One-Sample and Two-Sample Interlaboratory Testing Programs*, Annual Book of ASTM Standards, ASTM International, West Conshohocken, PA, 2016

Author: Paul Stutzman [paul.stutzman@nist.gov](mailto:paul.stutzman@nist.gov)  
Engineering Laboratory  
National Institute of Standards and Technology